

SoLLs.INTEC.13: International Conference on Knowledge-Innovation-Excellence: Synergy in Language Research and Practice

How do Arab speakers cognitively process Malay in the Arabic script?

Khazriyati Salehuddin^{a*}, Heather Winskel^b

^a*Faculty of Social Sciences & Humanities, Universiti Kebangsaan Malaysia, UKM Bangi, 43650, Malaysia*

^b*School of Health & Human Sciences, Southern Cross University, Coffs Harbour, NSW 2450, Australia*

Abstract

Reading is one area within cognitive psychology that is rigorously studied through experiments. Cognitive psychologists describe reading as “a highly complex process” because this receptive skill depends on different sub-processes that are inter-connected with one another. While studies on reading the Arabic script in the Arabic language have been (though not thoroughly) conducted before, a study on reading the Arabic script in non-Arabic languages by native and first language speakers of Arabic has never been conducted. This paper presents results of an experiment that investigated the reading of Malay in the Arabic script by Arabic native and first language speakers.

© 2013 The Authors. Published by Elsevier Ltd.

Selection and peer-review under responsibility of Universiti Kebangsaan Malaysia.

Keywords: Arabic script; cognitive complexity; DMDX; Malay; reading

1. Introduction

One of the most investigated areas within cognitive psychology is word reading (Plaut, 1997). Many of these studies are conducted from different perspectives (Rayner et. Al., 2012). They either investigate perceptual processing, or memory processing, or comprehension processing, or production processing. The difference in research focus on the different aspects of reading conducted thus far suggests that reading relies on various sub-processes that are inter-related to one another. This suggests reading as “a highly complex process” (Rayner et al., 2012, p.7).

Typically, experiments are used by cognitive psychologists to investigate the processing mechanisms involved in reading; this results in the “empirical experimentation” nature of methodology in the research conducted by

* Corresponding author. Tel.: +6-019-334-8256; fax: +6-03-8925-4577.

E-mail address: khazdin@ukm.my

cognitive psychologists (Rayner et al., 2012, p. 8). To date, various studies on reading have shown that frequency of occurrence, age of acquisition, orthographic length, regularity and consistency in spelling, and semantic characteristics of the words are among the various variables that influence the accuracy and speed of word reading (Cortese & Balota, 2012). Most reading research on skilled readers focuses on how words are recognized; hence the term word-recognition research. As a result, various word-recognition models have been introduced. Despite the difference in the models, these reading models revolve around investigating whether:

- word recognition begins by accessing whole-word representations in the mental lexicon or whether word recognition begins by accessing sub-word representations such as features, letters, or syllables;
- word identification occurs through “direct access to their meanings” or whether word identification occurs through “phonologically mediated access to word meanings”;
- letter processing in word recognition takes place in a serial manner or whether letter processing in word recognition takes place in a parallel/simultaneous manner;
- word recognition is primarily a context-driven (top-down) process or whether word recognition is primarily a stimulus-driven (bottom-up) process, or both (interactive);
- word recognition involves a single mechanism for accessing the lexicon or whether word recognition involves multiple mechanisms for accessing the lexicon;
- word activation does take place in word recognition or whether “search process” does take place in word recognition;

(Wolf , et.al., 1998)

Despite the various reading models introduced, letter discrimination is said to be the earliest stage of reading where novice readers “build up a repertoire of letter-sound associations that provides the basis for their move from being a pre-reader to reader” (Wolf, et. Al., 1998, p. 423). This is followed by the logographic phase, a phase where novice readers only recognize familiar words and are unable to read novel words. Novice readers then proceed to the alphabetic phase, where the grapheme-phoneme rule is applied before they proceed to the orthographic phase, a phase that involves the integration of phonological and morphological knowledge – with greater emphasis on comprehension rather than word identification per se.

Thus far, most of the reading research conducted have given emphasis on an array of existing writing system of various languages – alphabetic (e.g., Roman, Greek, Hebrew and Thai), syllabic (e.g., Japanese), and logographic (Chinese). However, at least to the researchers’ knowledge, despite being an alphabetic script, not much research has been conducted on the Arabic script, particularly on word recognition. Studies conducted by Abu-Rabia and colleagues (e.g., 2002, 2003) and Taouk and Coltheart (2004) are among the very few that looked into the processes involved in reading the Arabic script. Abu-Rabia and colleagues, for example, studied the features of the Arabic scripts whereas Taouk and Coltheart (2004) focused on the developmental stages involved in reading the script.

According to Taouk and Coltheart (2004), the Arabic writing system is learned in phases, i.e., the “discrimination-net” phase, followed by a phonological-recoding phase, and finally the orthographic phase. At the “discrimination-net” phase, novice (younger) readers do not use knowledge about relationships between letters and sound. They tend to generalize letters they see in one particular word to other words. For example, children who know “moon” will read the word “booth” as “moon” based on the double “o” (i.e., “oo”) they see in the word “booth”. Subsequently, at the phonological recoding phase, children apply their “grapheme-phoneme” rules to convert print into speech. At this phase, children should be able to translate visually unfamiliar strings of letters from print to speech. Hence, it is at this phase where children start to be able to read non-words (i.e., words that they have never read or those that do not exist in their language) correctly as if these words exist in their language (Taouk & Coltheart, 2004). For example, readers reading English who are at the phonological recoding stage should be able to read “blant” by translating each of the symbols into sounds, even though “blant” is not an

English word. The next phase is the orthographic phase; the phase where readers are known as skilled readers. This is the phase where readers rely on the context and on the spellings of words to decode a word. If the readers were readers reading English, these readers will exhibit their “uneasiness” when they are asked to read “blant” as a word due to their awareness of the fact that such a word does not exist in the English language.

The Arabic script is unique in its own ways. It is an alphabetic script and is written from right to left based on the “root and pattern” principles (i.e., the root of words represent the basic sense of words). The root of these words are formed by a string of consonants, and vowels are usually used to indicate “specific lexical meaning and also grammatical and syntactic information” (Abu-Rabia et al. 2003, p. 424). In the written form of the Arabic language, the vowels are sometimes manifested in the form of diacritics and for more advanced readers, the vowels are not manifested. Hence, despite the fact that the sound-symbol correspondence between the letters and the sound they represents is predictable, the correspondence is considered shallow only when the script is written with the vowel diacritics (or vowelized). For advanced readers, texts are orthographically “deep” (unvowelised, without diacritics) and readers have to include their knowledge of syntax, vocabulary, and contextual interpretation to understand the meaning.

While the above-mentioned studies conducted on Arabic scripts were on readers reading the Arabic language, could a similar description be made on reading the Arabic script in a language other than the Arabic language? It is a well-known fact that the Arabic script is not unique to Arabic. Other than Arabic, the Arabic script is used by other languages, namely Persian, Ottoman, Sindhi, Urdu, and Malay. Despite the inclusion of non-Arabic phonemes such as /p/ and /g/ in the Arabic script of these languages, the scripts are still considered as the Arabic script because of the same features that these derived Arabic scripts have with the original Arabic letters in terms of their forms (Abu-Rabia, 2002). For example, the letter that represents the sounds /g/ in the Malay language is represented with a symbol that has a close resemblance with the letter that represents the sound /k/ in Arabic.

Where the Arabic script for the Malay language is concerned, in terms of dominance, the Arabic script is the less dominant script in Malaysia in comparison to the Romanised script. Salehuddin (2012) suggests that the usage of the Arabic script is not as widespread as the Romanised script in Malay and is used particularly for religious purposes. This is due to the fact that reading the Arabic script is cognitively more complex for Malay native speakers in comparison to the Romanised script, particularly because of its physical properties. As put forward by earlier researchers (e.g., Abu-Rabia and colleagues), the physical features of some of the Arabic letters are strikingly similar to one another except for the number and the positioning of the “dots” that accompany the letters. For example, the phonemes /b/, /t/ and /θ/ share a similar shape except for the fact that /b/ has one dot underneath its body, /t/ has two dots above its body, whereas /θ/ has three dots above its body.

This paper presents a study carried out to investigate the reading performance of readers who are either native or first language speakers of Arabic in a rapid naming task. The words presented to them were words in the Malay language that are printed in the Arabic script. This research is conducted based on the assumption that readers who are native or first language speakers of Arabic will not have much difficulty in reading Malay in the Arabic script. It was hypothesized that the use of the Arabic script in the Malay language will facilitate reading in the Malay language due to the large similarity in the script of the two languages in terms of its form (shape of individual letters), written orientation (right-to-left for Arabic script), and presentation (cursive for Arabic).

2. Method

Ten readers who speak Arabic as a native or as a first language participated in this study. They were postgraduate students at Universiti Kebangsaan Malaysia (UKM) and regarded the Malay language as a foreign language. All participants had normal or corrected-to-normal vision. They were recruited based on the invitation circulated to several postgraduate students at the School of Language Studies and Linguistics, UKM. Prior to the session, each participant answered a questionnaire designed to gauge their fluency in reading Arabic, with and without vowel diacritics. A brief summary of their profile is as in Table 1.

Table 1. Participants' profile

	Gender		Reading Preference (Diacritics)		Reading speed (sec.)								
	Male	Female	With	Without	With				Without				
					2.0-3.0	3.0-4.0	4.0-5.0	5.0-6.0	1.0-2.0	2.0-3.0	3.0-4.0	4.0-5.0	5.0-6.0
Nos.	7	3	4	6	5	4	1	0	2	2	5	0	1

All participants were tested individually. Each of them sat approximately 50 cm in front of a *Multi-touch Full HD All-in-One computer*. Prior to the experimental session, each participant was briefed that Malay words written in Arabic scripts (printed in cursive) would be presented on the computer monitor and their task was to read the words aloud. Eighteen (18) bi-syllabic Malay words were presented to the participants as practice trails using the DMDX software and once they have read the 18 trials, they were asked to proceed to the experimental trials by pressing the “space bar” key. All instructions were given to the participants in English.

A total of 200 bi-syllabic Malay words were presented to the participants as experimental trials; 100 were bi-syllabic Malay words written without vowel diacritics whereas the other 100 were the same bi-syllabic Malay words but with vowel diacritics. The presentation of Malay words in the Arabic script with diacritics was based on the proposition put forward by Salehuddin (2013). The presentation of stimuli was randomized and mixed for both with and without diacritics, with different order for each participant. After the 100th stimulus, each participant was given a break and was instructed to resume the experimental trial by pressing the “space bar” key. The experimental session lasted for about 30 minutes/participant.

3. Results

An independent-samples t-test was conducted to compare the correct responses in reading Malay in Arabic scripts with diacritics and without diacritics. There was a significant difference in the scores for reading Malay in the Arabic script without diacritics condition ($M=69.1\%$, $S=6.74$) and reading Malay in the Arabic script with diacritics condition ($M=84.70\%$, $S=5.539$), $t(9)=6.642$, $p<.001$.

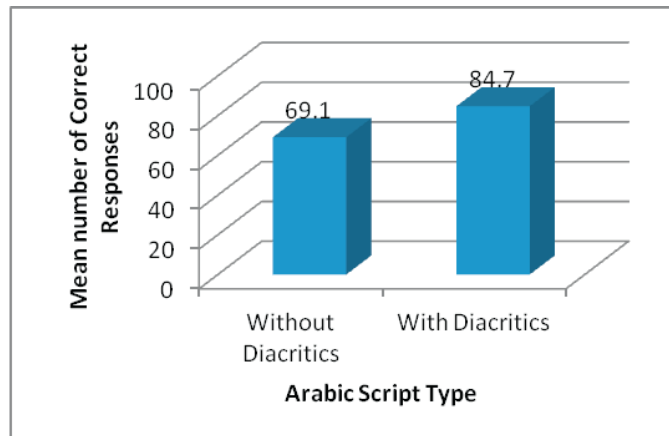


Figure. 1. Mean number of correct responses by Arabic native and first language speakers

4. Discussion and Conclusion

Despite the striking similarities in the form, orientation, and presentation of the Malay words presented in the experiment with the Arabic words, the performance of the Arabic native and first language speakers suggests that in reading Malay in the Arabic script, the Arabic native and first language speakers did face difficulty when reading Malay words in Arabic script. Difficulty in reading was more obvious when the Malay words were presented without diacritics in comparison to those presented with diacritics. Hypothetically, reading Malay in the Arabic script should not pose a problem to Arab native and first language speakers since Arabic vowel sounds are not manifested in the Arabic writing system. For example, the vowel /a/ is not overtly manifested in the Arabic script because despite the fact that the letter “*alif*” exists in the Arabic script for the Arabic language, the letter “*alif*” is not used to represent the short vowel /a/. Instead, “*alif*” is present in the Arabic script of the Arabic language to represent the low front long vowel /a:/.

The absence of diacritics in the Malay script in Arabic makes it difficult for Arabic native and first language speakers to read the words accurately, despite the presence of the vowel letters. This is possibly due to the fact that the manifestations of vowel phonemes in Malay syllables in the Arabic script are often done haphazardly (Salehuddin 2012). A syllable that contains the vowel /a/, for example, can be manifested in the Arabic script in two forms; the first with the vowel letter “*alif*” (ا), and the second without any vowel letter. When the first takes place, the script can be considered as “shallow”, because the presence of the letter “*alif*” can signal readers that the vowel /a/ should be manifested out loud in reading the syllable. However, when the second takes place, the script can be considered as “deep”, because the absence of the letter “*alif*” forces readers to guess the manner the syllable should be read. The Arabic script for the Malay language is considered to be cognitively complex because of its inconsistency in terms of whether it is a “shallow” orthography or whether it is a “deep” orthography within the same reading material. To illustrate, vowel /a/ in the Malay word “*mama*” is manifested in the form of the letter “*alif*” in both syllables; yet, the letter “*alif*” is not manifested in the first syllable of “*tajwid*”, in the second syllable of “*mata*”, and in neither syllables of “*tamar*”.

However, when vowels are manifested in the form of diacritics in the Arabic script of Malay, reading the Malay words appears to be a more “guided” reading task for the Arabic native and first language speakers. The presence of vowel diacritics helps them to decide which vowel phoneme should be used to make up a particular syllable; particularly when the words are not familiar to them. While all the native and first language speakers of Arabic indicated that they had no problem reading Arabic scripts without diacritics, their incorrect responses in the reading aloud task for Malay words without diacritics indicated that they did have problems in deciding which vowels should be used when reading the Malay words.

Why did the Arabic native and first language speakers not have problems when reading Arabic scripts in Arabic? As mentioned earlier, more advanced speakers of Arabic are often presented with the deep orthography of the Arabic script. This is because their experience in reading Arabic words and their knowledge with regard to the Arabic lexicon allows them to have a quick access to the words stored in their mental lexicon and hence, decipher the meaning of words and pronounce them correctly without having to read the individual letters that make up the word. Readers at this phase would almost immediately recognize particular words by accessing the whole-word representations in their mental lexicon. Hence, when advanced Arabic readers read a particular word, they would read the word as a whole, and possibly use the context where the word occurs to manifest the word phonologically (Wolf, et. Al., 1998).

A different situation takes place when Arabic native and first language speakers read Malay words that are written in the Arabic script. These readers who regard Malay as a foreign language naturally have limited access to the Malay words in their mental lexicon, and the amount of Malay words that are stored in their mental lexicon is typically small. Because of this, readers have to rely on all the visual clues that are available in the reading material when reading. Reading, for these readers, hence, had to be done in a bottom-up process, where words are recognized by accessing sub-word representations (i.e., letters) to make sense of the word. Due to the fact that

such words are not represented in their lexicon, particularly in the form of Arabic scripts, readers have to rely on whatever clues that are available in the text to help them pronounce the words they see. Because of this, the presence of the vowel diacritics would help these foreign language speakers of Malay to read Malay words that are unfamiliar to Arab native and first language speakers accurately, despite the fact that they have no knowledge on the semantics of the words and how they should be pronounced phonologically.

Interestingly, despite the fact that more Malay words are read more accurately when vowel diacritics are present in the words, there are some words (e.g., *alim*, *waqaf*, and *sifat*) that were read inaccurately despite the presence of the vowel diacritics. When analysed, these words are actually words that are borrowed from the Arabic language. It appears that similarity in the vocabulary between Arabic and Malay does not necessarily facilitate accuracy in the reading of Malay words (Lado, 1964). In fact, similarity between the two languages results in a greater interference to the Arabic native and first language speakers when reading the Malay words in the Arabic script. This supports the notion that word recognition of Malay words in the Arabic script, where Arabic native and first language speakers are concerned, is an interactive process. The bottom-up process is more dominant when the words are unfamiliar to the native and first language Arabic speakers. They were able to read Malay correctly when the words are presented with diacritics because they are at the letter discrimination phase – the earliest reading phase where novice readers “build up a repertoire of letter-sound associations that provides the basis for their move from being a pre-reader to reader” (Wolf, et.al., 1998, p. 423), a phase that is also known as the phonological recoding phase (Taouk & Coltheart, 2004). However, when the words appear to be familiar to the native and first language Arabic speakers because of the fact that the words are borrowed from the Arabic language, the role that vowel diacritics play in facilitating reading becomes minimal. This is due to the fact that for these words, native and first language Arabic speakers read them as they are at the orthographic phase; a phase that involves the integration of phonological and morphological knowledge that gives a greater emphasis on the word comprehension rather than word identification per se.

In conclusion, this study has shown that the cognitive processes that are involved in processing Malay words written in Arabic script are varied. The processes are not truly similar to the processes that are involved in reading Arabic words in Arabic script. The processes are dependent on various variables that lead to an interesting discussion on the reading processes. Nevertheless, to a certain extent, the presence of diacritics in the Malay scripts in Arabic does help native and first language speakers to pronounce Malay words more accurately.

Acknowledgements

The authors would like to thank the Ministry of Higher Education of Malaysia, for funding the ERGS/1/2011/SSI/UKM/02/5 research entitled *Exploring the Cognitive and Perceptual Processes in Reading*. Special gratitude also goes to all 10 postgraduate Arabic native and first language speakers of Universiti Kebangsaan Malaysia who participated in the study.

References

- Abu-Rabia, S. (2002). Reading in a root-based-morphology language: the case of Arabic. *Journal of Research in Reading*, 25(3), 299-309.
- Abu-Rabia, S., & Siegel, L. S. (2003). Reading skills in three orthographies: The case of trilingual Arabic-Hebrew-English-speaking Arab children. *Reading and writing*, 16(7), 611-634.
- Cortese, M. J., & Balota, D. A. (2012). Visual Word Recognition in Skilled Adult Readers. *The Cambridge Handbook of Psycholinguistics*, 159. James, C. *Contrastive Analysis*, Essex: Longman; 1980.
- Lado, R. (1964). *Language Teaching: A Scientific Approach*, New York: McGraw-Hill.

- Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and cognitive processes*, 12(5-6), 765-806.
- Rayner, K, Pollatsek, A, Ashby, J & Clifton, C. (2012). *Psychology of Reading*, New York: Psychology Press.
- Salehuddin, K. (2012). Penilaian ke atas kerumitan kognitif dalam proses membaca jawi. *GEMA: Online Journal of Language Studies*, 12(4), 1179-1192.
- Salehuddin, K. (2013). Arabic script of written Malay: Innovative transformations towards a less complex reading process. *Pertanika Journal of Social Sciences and Humanities*; 21(S): 61-74.
- Taouka, M., & Coltheart, M. (2004). The cognitive processes involved in learning to read in Arabic. *Reading and writing*, 17(1-2), 27-57.
- Wolf, M., Vellutino, F., & Gleason, J. B. (1998). A psycholinguistic account of reading. *Psycholinguistics*, 2, 409-451.